PERSPECTIVE: ANALYTICAL BIOTECHNOLOGY

Protein Structure Determination in Solution by Two-Dimensional and Three-Dimensional Nuclear Magnetic Resonance Spectroscopy

Angela M. Gronenborn* and G. Marius Clore*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Over the last decade, nuclear magnetic resonance (NMR) spectroscopy has evolved into a powerful method for determining structures of biological macromolecules. This has opened a unique opportunity for obtaining high-resolution three-dimensional structures in *solution*, in contrast to the well-established methods of X-ray diffraction, which are applicable only to solids and in particular single crystals. This rapid development has been spurred by several key advances in the field, especially the introduction of two- and three-dimensional NMR experiments, high field spectrometers (500 and 600 MHz), and computational algorithms for converting NMR derived restraints into three-dimensional structures. This review outlines the methodology employed for solving protein structures in solution, describing the basic NMR experiments necessary as well as introducing the concepts upon which the computational algorithms are founded. A variety of examples is discussed, illustrating the present state of the art, and future possibilities are indicated.

INTRODUCTION

Over the last few years there has been a burst of renewed interest in protein studies, especially aimed at understanding their structures, functions, and physiological roles. Some of this increased enthusiasm can be attributed to a variety of technological advances in the area of modern molecular biology. The advent of molecular cloning, particularily of protein encoding genes, and rapid DNA sequencing methods have led to an explosion in the number of available protein sequences, reaching over 15 000 at the latest count. Whereas this wealth of data is clearly impressive, amino acid sequences per se are of limited value in understanding protein function. We need to know the three-dimensional structure before we can begin to make progress in analyzing the intricate reactions carried out by proteins such as catalysis, ligand binding, gene regulation, and assembly. The design of modified proteins and rational design of ligands (drugs), as well as attempts at de novo design of protein molecules, all have to be based on concepts at the atomic level in three-dimensional space, thereby creating an increasing need for detailed structural analysis.

Until recently, the only experimental technique available for determining three-dimensional structures had been single-crystal X-ray diffraction, and most of our structural knowledge about proteins is based on those crystal structures. There are approximately 400 coordinate sets available to date, comprising about 120 different protein folds. Analyzing protein structures by crystallography can be a slow and difficult undertaking since the first, and possibly hardest task, involves growing X-ray quality grade crystals, which have to be well ordered to give rise to good diffraction spots. Even if this task is accomplished, a second hurdle still needs to be overcome, as the phases have to be solved, commonly achieved by collecting data on heavy atom derivatives. Thus, despite spectacular advances in protein crystallography, we are faced with an enormous gap between the available primary sequence data and the tertiary structure data on proteins.

Over the last ten years, a second method for determining protein structures has been developed and is by now well established. This method makes use of nuclear magnetic resonance (NMR) spectroscopy. Unlike crystallography, NMR measurements are carried out in solution under potentially physiological conditions and are therefore not hampered by the ability or inability of a protein to crystallize.

The principal source of information used to solve three-dimensional protein structures by NMR spectroscopy resides in short interproton distances supplemented by torsion angles. The distances are derived from nuclear Overhauser effect (NOE) measurements. The size of the NOE between two protons is proportional to r^{-6}, where r is the distance between them. Torsion angles are obtained from an analysis of three-bond coupling constants, which are related to dihedral angles. An essential prerequisite for obtaining interproton distance restraints and torsion angle restraints is the assignment of the NMR spectrum; that is to say the identity of every proton resonance has to be determined. This is not a trivial task considering that the proton spectrum of even a small protein comprising only 80 amino acids contains approximately 650 resonances. All of these exhibit several cross-peaks in the two-dimensional (2D) spectrum such that the number of cross-peaks can easily reach several thousand whose identity has to be ascertained. Complete spectral assignment is therefore an integral part of the structure determination.

Although it was appreciated relatively early on that NMR could in principle provide the necessary information to obtain three-dimensional structures, it is only fairly recently that this goal has been realized. The reasons for this are 3-fold: (i) The development of 2D NMR experiments (*1–3*) alleviated problems associated with resonance overlap, which for macromolecules prevents any analysis of the traditional one-dimensional spectrum. This is achieved by spreading all the information out in a plane, thereby permitting a detailed interpretation of the pertinent spectral features. This conceptual idea has been extended more recently to three-dimensional (3D) NMR (*4–7*) again relieving problems arising

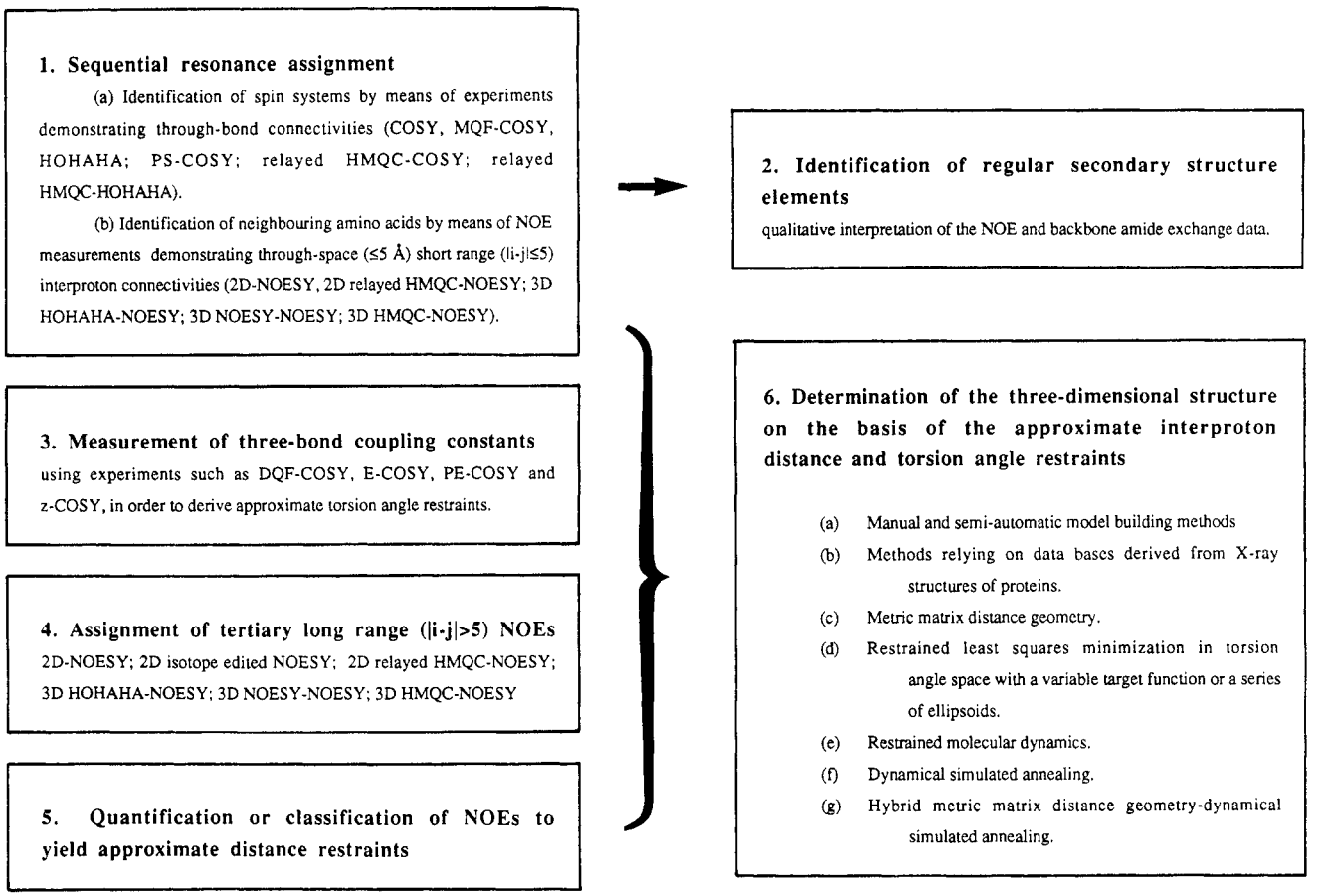


Figure 1. Flow chart of the various steps involved in determining the three-dimensional structure of a protein in solution by NMR.

from spectral crowding, and it is the 3D approach in particular that will extend the present limits with respect to the size of the proteins that can be studied. (ii) The availability of high field magnets (500 and 600 MHz) has resulted in spectrometers with both a significant increase in signal-to-noise ratio and greater spectral resolution, and continuing development in this area again will extend the limits even further. (iii) Suitable mathematical algorithms and computational approaches that convert the NMR derived restraints into three-dimensional structures have been developed (*8–14*) and several robust and efficient methods are now available.

This article reviews the various stages involved in the determination of a three-dimensional protein structure by NMR, and the flow chart in Figure 1 illustrates the individual steps. The general methodology is outlined; however, no attempt is made to provide an in-depth description of either the general NMR theory or the details of the mathematical algorithms. Emphasis is placed on the application of NMR to structural studies and several examples illustrating various points are presented.

BASIS OF TWO- AND THREE-DIMENSIONAL NMR

The principles of 2D NMR have been discussed in depth (see ref 15 for a comprehensive review) and only a very basic and brief description will be given here. Each proton (spin) possesses a property known as magnetization. When a molecule is placed in a magnetic field, the magnetization lies parallel to it. Rotation of this magnetization away from its parallel orientation, either by a radio frequency pulse or a combination of pulses allows one to follow the return of the magnetization to its equilibrium state. In any 2D NMR experiment this is called the preparation period. This is followed by an evolution period in which this transient state of the spins is allowed to evolve for varying time periods t_1, a mixing period

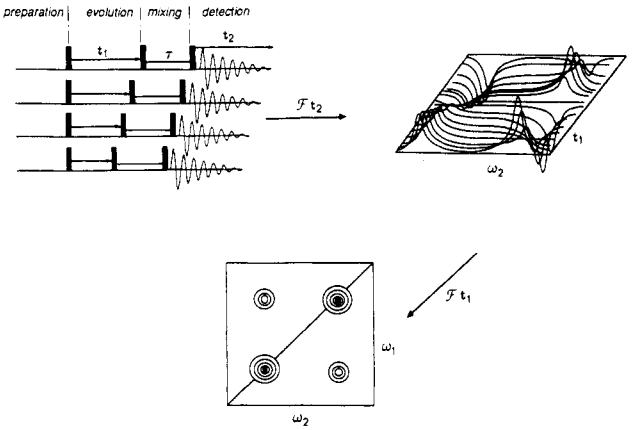


Figure 2. Generalized representation of the 2D NMR experiment.

during which the spins are correlated with each other, and finally the detection period t_2. A number of experiments are recorded with increasing values for t_1 to generate a data matrix $s(t_1,t_2)$. Two-dimensional Fourier transformation of $s(t_1,t_2)$ yields the 2D spectrum $S(\omega_1,\omega_2)$. The second frequency dimension in the 2D spectrum originates from the Fourier transformation of the t_1 modulation paptterns in the 1D spectra. The two frequency coordinates ω_1 and ω_2 of a particular resonance thus correspond to the t_1 and t_2 frequencies associated with the observed magnetization. This is illustrated schematically in Figure 2. Most homonuclear 2D experiments contain the 1D spectrum on the diagonal with symmetrically placed cross-peaks on either side of the diagonal representing different kinds of interactions between the spins. The nature of the interaction depends on the type of experiment, with cross peaks arising from through-bond scalar interactions in a COSY (*co*rrelated *s*pectroscopy) (*2*) or HOHAHA (*homo*-

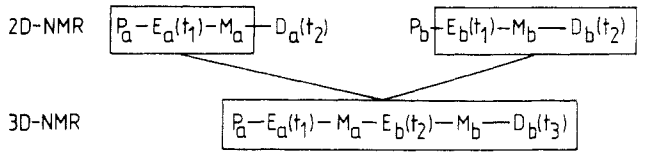


Figure 3. Scheme for combining two 2D NMR experiments into a 3D NMR experiment.

nuclear *Ha*rtmann–*Ha*hn) (*16*) experiment and from through-space correlations in a NOESY (*n*uclear *O*verhauser and *e*xchange *s*pectroscopy) (*17*) experiment.

Since every 2D NMR experiment consists of the basic scheme

preparation–evolution (t_1)–mixing–detection (t_2)

it is conceptually very simple to construct a 3D experiment out of two 2D experiments by omitting the detection period of the first 2D experiment and the preparation period of the second one, combining both into a single pulse train as illustrated in Figure 3. The new detection period is now called t_3 and for every time variable t_2, a complete (t_1,t_3) 2D data set is acquired. Fourier transformation of t_2 sections taken through these 2D data sets converts the set of 2D spectra into the 3D spectrum. Since any two 2D experiments can be combined into a 3D experiment, one can envisage an enormous number of such experiments, starting from homonuclear versions and progressing to various heteronuclear 3D experiments. A recent overview over a large number of combinations of 2D experiments for 3D versions can be consulted for further reading (*18*). The power of the 3D NMR experiments lies in overcoming resonance overlap which for larger proteins becomes again a major obstacle in the traditional 2D ones. The first 3D experiments on proteins were of the homonuclear type (*4*, *5*). While elegant and no doubt useful in certain cases, the applicability of these homonuclear 3D experiments to larger proteins is limited, as the efficiency of magnetization transfer is severely reduced with increasing line widths. This is not a problem for 3D heteronuclear experiments (*6*, *7*) which contain a heteronuclear shift correlation (HMQC, *h*eteronuclear *m*ultiple *q*uantum *c*oherence) experiment such as 3D NOESY-HMQC or 3D HOHAHA-HMQC. For those the heteronuclear magnetization transfer occurs via relatively large one-bond couplings and is therefore very effective. Such heteronuclear 3D experiments represent in essence a series of 2D NOESY or HOHAHA spectra edited with respect to the chemical shift of the directly bonded heteronucleus, such as ^{15}N or ^{13}C. Because these spectra can be regarded as stretched out heteronuclear edited 2D spectra, their basic appearance and features resemble 2D spectra, which allows for easy interpretation and data analysis. A schematic drawing of such a 3D spectrum is presented in Figure 4. Major advantages of these heteronuclear 3D experiments are their high sensitivity (provided the protein is isotopically enriched) and ease of analysis.

NUCLEAR OVERHAUSER EFFECT

NMR-derived protein structures are mainly based on NOE measurements that can demonstrate the proximity of protons in space and allow determination of their approximate separation (*19–21*). The principle of the NOE is relatively straightforward and is summarized in Figure 5. Considering the simplest system with only two protons, each of which possesses a property known as magnetization, exchange of magnetization between the protons occurs by a process known as cross-relaxation. Because the cross-relaxation rates in both directions are equal, the magnetization of the two protons at equilibrium is equal. The approximate chemical analogy of such a system would be one with two interconverting species with an equilibrium constant of 1. The cross-relaxation rate

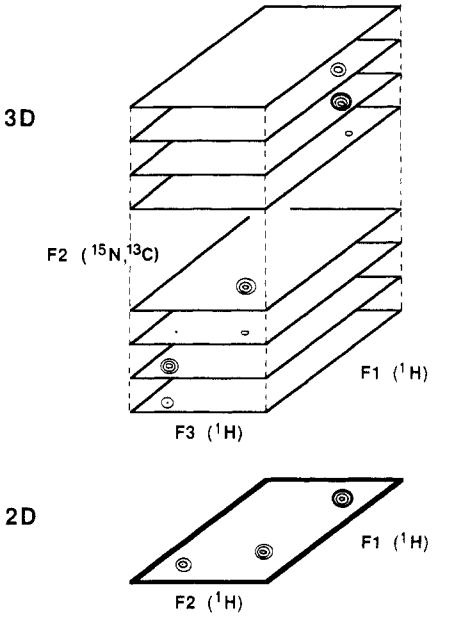
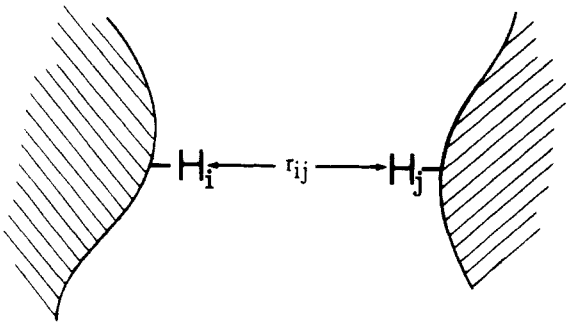


Figure 4. Schematic illustration of a heteronuclear 3D NMR experiment.



$$N_{ij}(t) \sim \sigma_{ij}\, t$$

$$\sigma_{ij} = \frac{\gamma^4 \hbar^2}{10\, r_{ij}^6}\left(\tau_{app} - \frac{6\tau_{app}}{1 + 4\omega^2\tau_{app}^2}\right)$$

Figure 5. Basis of the NOE: r_{ij}, distance between the protons i and j; σ, cross-relaxation rate; N, NOE; τ_{app}, correlation time.

is proportional to two variables: r^{-6}, where r is the distance between the two protons, and τ_{app} the effective correlation time of the interproton vector. It follows that if the magnetization of one of the spins is perturbed, the magnetization of the second spin will change. In the case of macromolecules the cross-relaxation rates are positive and the leakage rate from the system is very small, so that, in the limit, the magnetization of the two protons would be equalized. The change in magnetization of proton i upon perturbation of the magnetization of proton j is known as the nuclear Overhauser effect (NOE). The initial build-up rate of the NOE is equal to the cross-relaxation rate and, hence, proportional to r^{-6}.

In one-dimensional NMR, the NOE can be observed in a number of ways, all of which involve the application of a selective radio frequency pulse at the position of one of the resonances. The simplest experiment involves the irradiation of resonance i for a time t, followed by acquisition of the

spctrum. If proton j is close in space to proton i, its magnetization will be reduced and this is best observed in a difference spectrum subtracting a spectrum without irradiation from one with selective irradiation. An alternative approach involves the selective inversion of resonance *i* followed by acquisition after a time *t*. This particular experiment is the one-dimensional analogue of the two-dimensional experiment. In the two-dimensional experiment cross-peaks between proton resonances *i* and *j* are observed when the two protons are close in space and thus exchange magnetization via cross-relaxation.

SEQUENTIAL RESONANCE ASSIGNMENT

Sequential resonance assignment of the ^{1}H NMR spectra of proteins relies on two sorts of experiments: (i) those demonstrating through-bond scalar connectivities, and (ii) those demonstrating through-space (<5 Å) connectivities (*22*). The former, which are generally referred to as correlation experiments, serve to group together protons belonging to the same residue. The latter involve the detection of NOEs and serve to connect one residue with its immediate neighbors in the linear sequence of amino acids.

The first step in the assignment procedure lies in identifying spin systems, that is to say protons belonging to one residue unit in the polypeptide chain. Such experiments have to be carried out both in H_2O and D_2O, the former to establish connectivities involving the exchangeable NH protons and the latter to identify connectivities between nonexchangeable protons. Some spin systems are characteristic of single amino acids. This is the case for Gly, Ala, Thr, Leu, Ile, and Lys. Others are characteristic of several different amino acids. For example, Asp, Asn, Cys, Ser, and the aliphatic protons of all aromatic amino acids belong to the AMX spin system (i.e. they all have one C$^\alpha$H and two C$^\beta$H protons).

The simplest experiment used to delineate spin systems via scalar correlations is the COSY experiment, which was first described in 1976 by Aue et al. (*2*) and demonstrates direct through-bond connectivities. Thus, for a residue which has NH, C$^\alpha$H, C$^\beta$H, and C$^\gamma$H protons, connectivities will only be manifested between the NH and C$^\alpha$H, C$^\alpha$H and C$^\beta$H, and C$^\beta$H and C$^\gamma$H protons. This basic COSY experiment has now been superseeded by slightly more sophisticated experiments such as DQF-COSY (*23*) and P.COSY (*24*), which have the advantage of exhibiting pure phase absorption diagonals when the spectra are recorded in the pure phase absorption mode. This enables one to detect cross-peaks close to the diagonal.

Experiments that demonstrate only direct through-bond connectivities are of limited value if taken alone, due to problems of spectral overlap. In a protein spectrum, the degree of spectral overlap tends to increase as one progresses from the NH and C$^\alpha$H protons to the side chain protons. For this reason, experiments that also demonstrate indirect or relayed through-bond connectivities, for example between the NH and C$^\beta$H protons, are invaluable. In this respect, the most useful experiment is the HOHAHA experiment (*16*) (also referred to as TOCSY for *t*otal *c*orrelated *s*pectroscopy (*25*)). By adjusting the experimental mixing time, one can obtain successively direct, single, double, and multiple relayed connectivities. Further, the multiplet components of the cross-peaks are all in-phase in HOHAHA spectra, in contrast to COSY type spectra where they are in antiphase. As a result, the HOHAHA experiment is in general more sensitive and affords better resolution than the COSY type experiment. A schematic representation of cross-peak patterns observed in HOHAHA spectra for the various spin systems is illustrated in Figure 6. Examples of protein HOHAHA spectra in H_2O and D_2O are shown in Figure 7.

Once a few spin systems have been identified, one can then proceed to identify sequential through-space connectivities

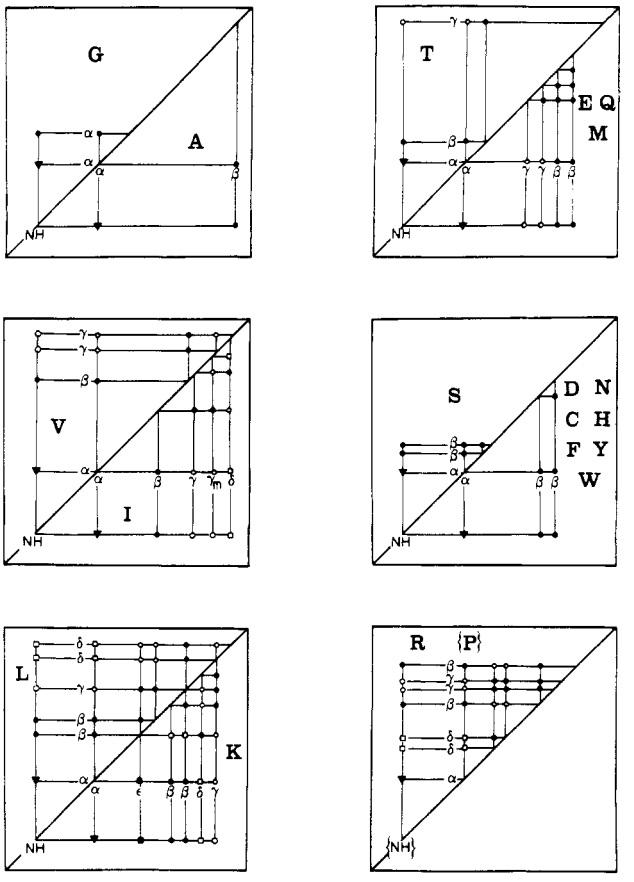


Figure 6. Schematic representation of cross-peak patterns observed in HOHAHA spectra for various amino acid spin systems. In a COSY spectrum, only direct connectivities are observed.

involving the NH, C$^\alpha$H, and C$^\beta$H protons by means of two-dimensional NOE spectroscopy (NOESY). For the purpose of sequential assignment the most important connectivities are the C$^\alpha$H(*i*)–NH (i + 1, 2, 3, 4), C$^\beta$H(*i*)–NH(i + 1), NH-(*i*)–NH(i + 1), and C$^\alpha$H(i)–C$^\beta$H(i + 3) NOEs. This is illustrated schematically in Figure 8 and an example of a NOESY spectrum is shown in Figure 9.

In the case of large molecules where the NH proton resonances are broad, the sensitivity of the conventional COSY spectrum can be improved by a factor of ~2 by recording a ^{15}N-filtered COSY spectrum (*26*) (also known as PS-COSY for *p*seudo *s*ingle quantum COSY). Labeling the protein uniformly with ^{15}N allows for efficient generation of heteronuclear zero and double quantum coherences whose relaxation rates, to a first-order approximation, are not affected by heteronuclear dipolar coupling. This permits one to eliminate one of the major line broadening mechanisms for amide protons in proteins, namely heteronuclear dipolar coupling to the nitrogen nucleus. As a result the multiple quantum resonances are significantly narrower than the corresponding NH resonances. The ^{15}N chemical shift contribution is easily removed from the multiple quantum frequency, yielding spectra that are similar in appearance to a regular COSY spectrum apart from the line narrowing of the NH resonances.

As proteins get larger, problems associated with chemical shift dispersion become increasingly severe. One approach for alleviating such problems in the sequential assignment of proteins involves correlating proton–proton through-space and through-bond connectivities with the chemical shift of a directly bonded NMR active nucleus such as ^{15}N or ^{13}C. In the case of completely ^{15}N labeled protein, two sorts of experiments are particularly useful. The first are relayed experiments combining the heteronuclear multiple quantum coherence scheme (*27–31*) with experiments such as NOESY,

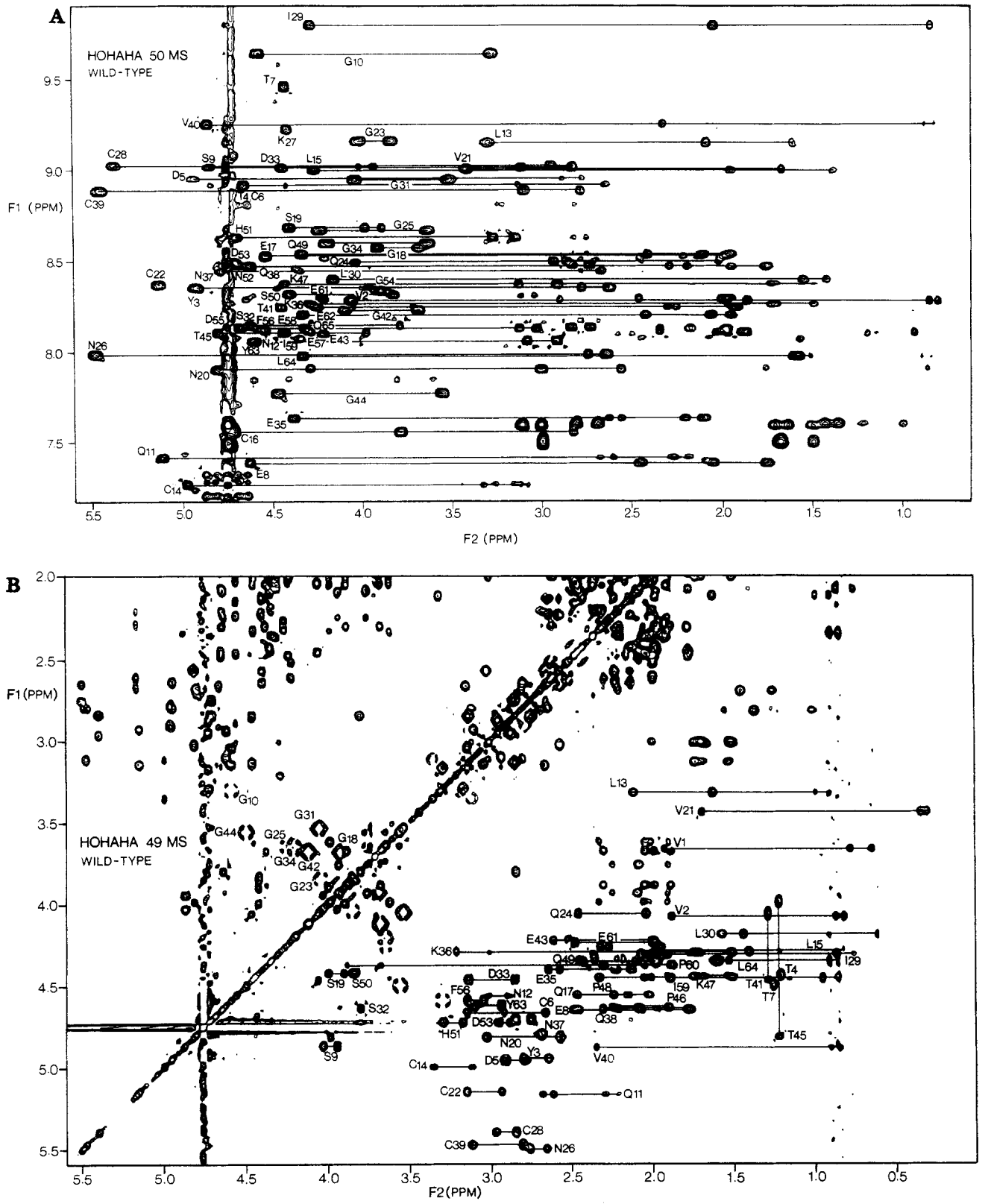


Figure 7. 600-MHz HOHAHA spectrum of hirudin in H_2O showing the NH(F2)–aliphatic(F1) region of the spectrum. Peaks are labeled at the positions of the direct NH–C$^{\alpha}$H connectivities (A). 600-MHz HOHAHA spectrum of hirudin in D_2O showing the C$^{\alpha}$H(F1)–aliphatic(F2) region of the spectrum. A number of spin systems are indicated (B) (ref 50).

COSY, or HOHAHA (Figure 10) (*32*). These experiments yield the same information as the homonuclear NOESY, COSY, and HOHAHA experiments but the NH proton chemical shift axis is replaced by that of the ^{15}N chemical shift. Because it is rare to find that both the ^{1}H and ^{15}N chemical shifts of two NH groups are degenerate, NOEs and through-bond correlations involving NH protons with the same chemical shift can readily be resolved in this manner. The second type of experiment involves the detection of long range correlations between ^{15}N and C$^{\alpha}$H atoms using ^{1}H-detected *h*eteronuclear *m*ultiple-*b*ond correlation (HMBC) spectroscopy (*33*). In particular, the observation of two-bond $^{15}N(i)$-C$^{\alpha}$H(i) and three-bond $^{15}N(i)$-C$^{\alpha}$H(i-1) correlations enables one to connect one residue with the next (*34*). Additionally, because

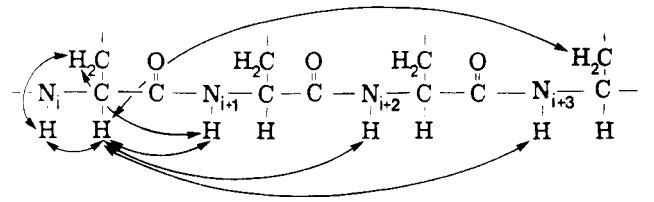


Figure 8. Schematic illustration of the connectivities used for sequential resonance assignment of protein spectra.

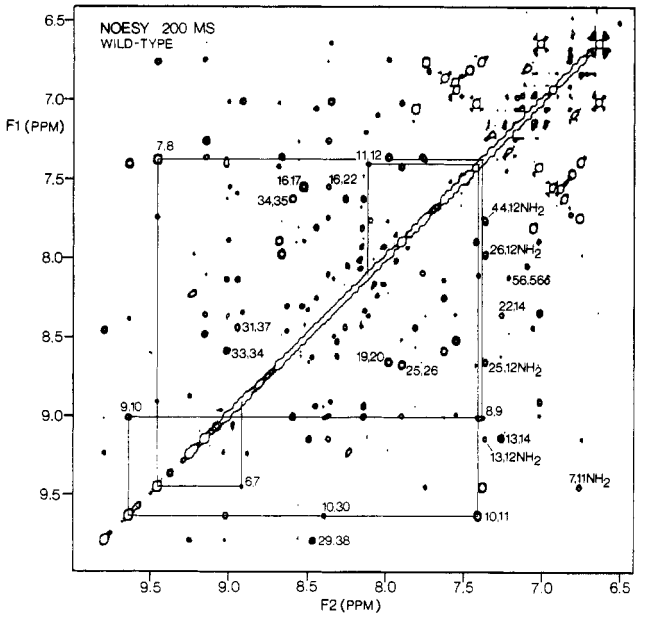


Figure 9. 600-MHz NOESY spectrum of hirudin in H_2O showing the NH(F1)–NH(F2) region of the spectrum. A stretch of NH(i)–NH(i + 1) connectivities extending from residues 6 to 12 is indicated. In addition several long-range NOEs are marked (ref 50).

the size of the $^{15}N(i)$–$C^{\alpha}H(i-1)$ coupling is very sensitive to the Ψ backbone torsion angle, qualitative structural information can readily be derived.

A further avenue for large proteins involves the application of three-dimensional NMR. For ^{15}N 3D spectra, the normal rules for making sequence specific assignments can be readily applied. The only difference to the normal 2D case is that connections between one residue and the next must be made not only between different sets of peaks but also between different planes of the spectrum. The major advantage going from 2D to 3D is that the distribution of overlapping or closely spaced cross-peaks throughout the entire cube removes most of the ambiguities present in the 2D spectrum. A comparison of the ^{15}N edited 2D spectra with slices from the corresponding 3D spectra is presented in Figure 11, demonstrating the dramatic reduction in the number of cross-peaks. Combining such slices from a 3D ^{15}N NOESY-HMQC spectrum with those of a ^{15}N HOHAHA-HMQC spectrum allows the straightforward sequential assignment by hopping from one pair of HOHAHA/NOESY planes to another pair, connecting them either via $C^{\alpha}H(i)$–NH(i + 1) or NH(i)–NH(i + 1) NOEs in the same manner as for the 2D case. This is illustrated for a stretch of residues for the protein interleukin-1β in Figure 12.

IDENTIFICATION OF REGULAR SECONDARY STRUCTURE ELEMENTS

Because each type of secondary structure element is characterized by a particular pattern of short range ($|i-j| \leq 5$) NOEs (*22*, *35*), qualitative interpretation of the sequential NOEs allows the identification of regular secondary structure elements. This is illustrated in Figure 13. Thus, for example, helices are characterized by a stretch of strong or medium NH(i)–NH(i + 1) NOEs and medium or weak $C^{\alpha}H(i)$–NH(i + 3) and $C^{\alpha}H(i)$ – $C^{\beta}H(i+3)$ NOEs and $C^{\alpha}H(i)$ – NH(i + 1) NOEs, sometimes supplemented by NH(i) – NH(i + 2) and $C^{\alpha}H(i)$ – NH(i + 4) NOEs. Strands, on the other hand, are characterized by very strong $C^{\alpha}H(i)$ – NH(i + 1) NOEs and by the absence of other short-range NOEs involving the NH and $C^{\alpha}H$ protons. β sheets can be identified and aligned from interstrand NOEs involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons. It should also be pointed out that the identification of secondary structure elements is aided by NH exchange data, in so far that slowly exchanging NH protons are usually involved in hydrogen bonding, and by $^3J_{HN\alpha}$ coupling constant data. Figure 14 illustrates the application of this method to the protein interleukin-8 (*36*). Inspection of the short range NOE data immediately enables one to identify three β strands, several turns, and a long helix at the carboxy terminus.

Several factors should be borne in mind in assessing the accuracy of secondary structure elements deduced by using this approach. Essentially it is a data based approach in so far that the expected patterns of short range NOE connectivities for different secondary structure elements have been derived by examining the values of all the short range distances involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons in regular secondary structure elements present in protein X-ray structures. Thus, it tends to perform relatively poorly in regions of irregular structure such as loops. In addition, the exact start and end of helices tend to be rather ill-defined, particularly as the pattern of NOEs for turns is not all too dissimilar from that present in helices. Thus, a turn at the end of a helix could be misinterpreted as still being part of the helix. In the case of β sheets, the definition of the start and end is more accurate as the alignment is accomplished from the interstrand NOEs involving the NH and $C^{\alpha}H$ protons. Therefore, although this secondary structure delineation is a very easy and straightforward procedure, it can only be used in a qualitative fashion and accurate positioning of the identified secondary structure elements can be only accomplished after the complete 3D protein structure has been determined.

INTERPROTON DISTANCES RESTRAINTS

The initial slope of the time-dependent NOE, $N_{ij}(t)$, between two protons i and j is equal to the cross-relaxation rate σ_{ij} between the two protons (*37*)

$$\left.\frac{dN_{ij}}{dt}\right|_{t=0} = \sigma_{ij} \tag{1}$$

σ_{ij} is simply the rate constant for exchange of magnetization between the two protons. σ_{ij}, in turn, is proportional to $\langle r_{ij}^{-6} \rangle$ and $\tau_{eff}(ij)$, where r_{ij} is the distance between the two protons and $\tau_{eff}(ij)$ the effective correlation time of the i–j vector:

$$\sigma_{ij} = \frac{\gamma^4 \hbar^2}{10 r_{ij}^6}\left(\tau_{eff}(ij) - \frac{6\tau_{eff}(ij)}{1 + 4\omega^2 \tau_{eff}(ij)^2}\right) \tag{2}$$

(γ is the gyromagnetic ratio of the proton, $\hbar$ is Planck's constant divided by 2π, and ω is the spectrometer frequency). It therefore follows that at short mixing times, τ_m, ratios of NOEs can yield either ratios of distances or actual distances, if one distance is already known, through the relationship

$$r_{ij}/r_{kl} = (\sigma_{kl}/\sigma_{ij})^{1/6} \tag{3}$$

providing the effective correlation times for the two interproton vectors are approximately the same.

In practice, initial slope measurements are not entirely trivial. First the magnitude of the NOEs at very short mixing times are small, inevitably posing a single-to-noise problem. Second, the measured NOE at short mixing times may not

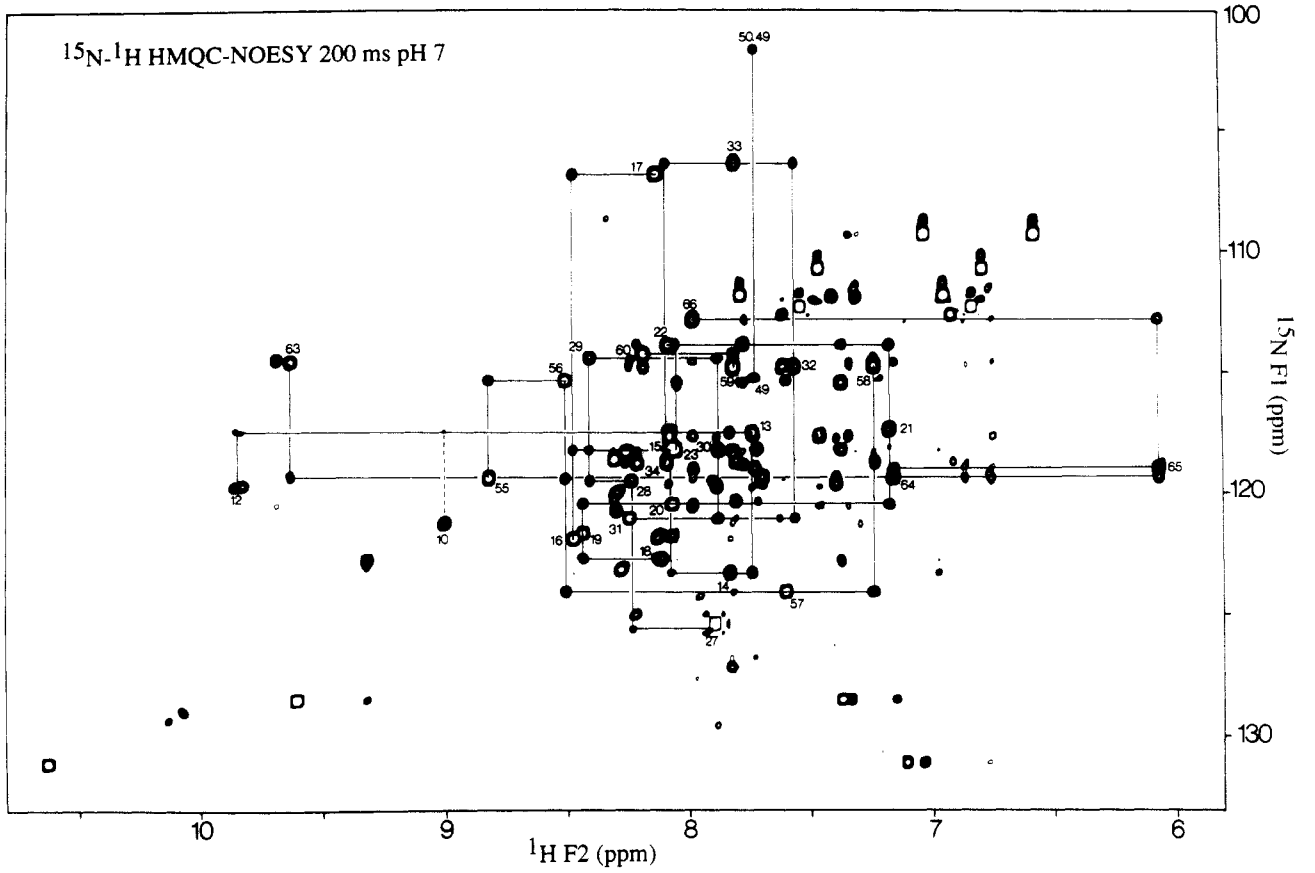


Figure 10. 600-MHz ^{15}N–^{1}H HMQC-NOESY spectrum of uniformly labeled ^{15}N Ner protein in H_2O. The ^{15}N(F1)–NH(F2) region of the spectrum with selected NH(i)–NH(i + 1) NOE connectivities is shown (ref 32).

reflect the true magnitude of the NOE due to the particularities of the experimental setup. In addition, problems arising from spin diffusion in large multispin systems introduce errors into the initial rate approximation. These can to some degree be minimized by carrying out a full relaxation matrix analysis (*38*), although certain problems still exist (*39*). For proteins, however, variations in effective correlation time will always present a problem if one wants to extract accurate interproton distances and it is therefore advisable to settle for only approximate interproton distance restraints. Because of the $\langle r^{-6} \rangle$ dependence of the NOE, such approximate interproton distance restraints can clearly be derived even in the presence of large variations in effective correlation times. Empirically, the type of classification generally used is one in which strong, medium, and weak NOEs correspond to distance ranges of approximately 1.8–2.7, 1.8–3.3, and 1.8–5.0 Å, where the lower limit of 1.8 Å corresponds to the sum of the van der Waals radii of two protons. By use of such a scheme, variations in effective correlations times do not introduce errors into the distance restraints. Rather, they only result in an increase in the estimated range for a particular interproton distance.

TORSION ANGLE RESTRAINTS

Vicinal spin–spin coupling constants can provide useful information supplementing the interproton distance restraints derived from NOE data. In particular, ranges of torsion angles can be estimated from the size of the coupling constants. The latter may be obtained by analyzing the multiplet patterns in COSY and COSY-like (e.g. DQF-COSY, E-COSY, PE-COSY, z-COSY) spectra.

The easiest coupling constants to determine in proteins are the $^{3}J_{HN\alpha}$ coupling constants that can be obtained by simply measuring the peak-to-peak separation of the antiphase components of the C$^{\alpha}$H–NH COSY cross-peaks. The size of the $^{3}J_{HN\alpha}$ coupling constant is related to the ϕ backbone torsion angle through a Karplus-type relationship (*40*). Consequently, values of $^{3}J_{HN\alpha}$ <6 and >8 Hz correspond to ranges of −10° to −90° and −80° to −180°, respectively, for the ϕ backbone torsion angles. Considerable care, however, has to be taken in deriving ϕ backbone torsion angle ranges from apparent values of $^{3}J_{HN\alpha}$ coupling constants measured in this way, as the minimum separation between the antiphase components of a COSY cross-peak is equal to approximately half of the NH line width (*41*). That is to say that small coupling constants can only be determined for relatively sharp resonances. This limitation can be overcome by measuring $^{3}J_{HN\alpha}$ from ^{15}N HMQC-COSY/HMQC-J spectra (*42*) because of the significantly narrower multiple quantum line widths in these experiments.

χ_1 side chain torsion angle restraints and stereospecific assignments can be obtained by analyzing the pattern of $^{3}J_{\alpha\beta}$ coupling constants and the relative intensities of the intraresidue NOEs from the NH and C$^{\alpha}$H protons on the one hand to the two C$^{\beta}$H protons on the other, and in the case of valine to the C$^{\gamma}$H$_3$ protons. The $^{3}J_{\alpha\beta}$ coupling constants are related to the χ_1 torsion angle and are best measured from correlation spectra which yield reduced multiplets such as β-COSY, E-COSY, P.E.COSY, or z-COSY (*43*, *44*). In addition, under suitable conditions, they can be qualitatively assessed from the C$^{\alpha}$H–C$^{\beta}$H cross-peak shapes in HOHAHA spectra (*45*).

If both $^{3}J_{\alpha\beta}$ couplings are small (~3 Hz) then χ_1 must lie in the range 60 ± 60°. If, on the other hand, one of the $^{3}J_{\alpha\beta}$ couplings is large and the other small, χ_1 can lie either in the range 180 ± 60° or −60 ± 60°. These two possibilities are easily distinguished on the basis of short mixing time NOESY experiments which yield simultaneously stereospecific assignments of the β-methylene protons. Clearly, this approach may fail if a side chain has a mixture of conformations or the χ_1 angle deviates by more than ~40° from the staggered rotamer

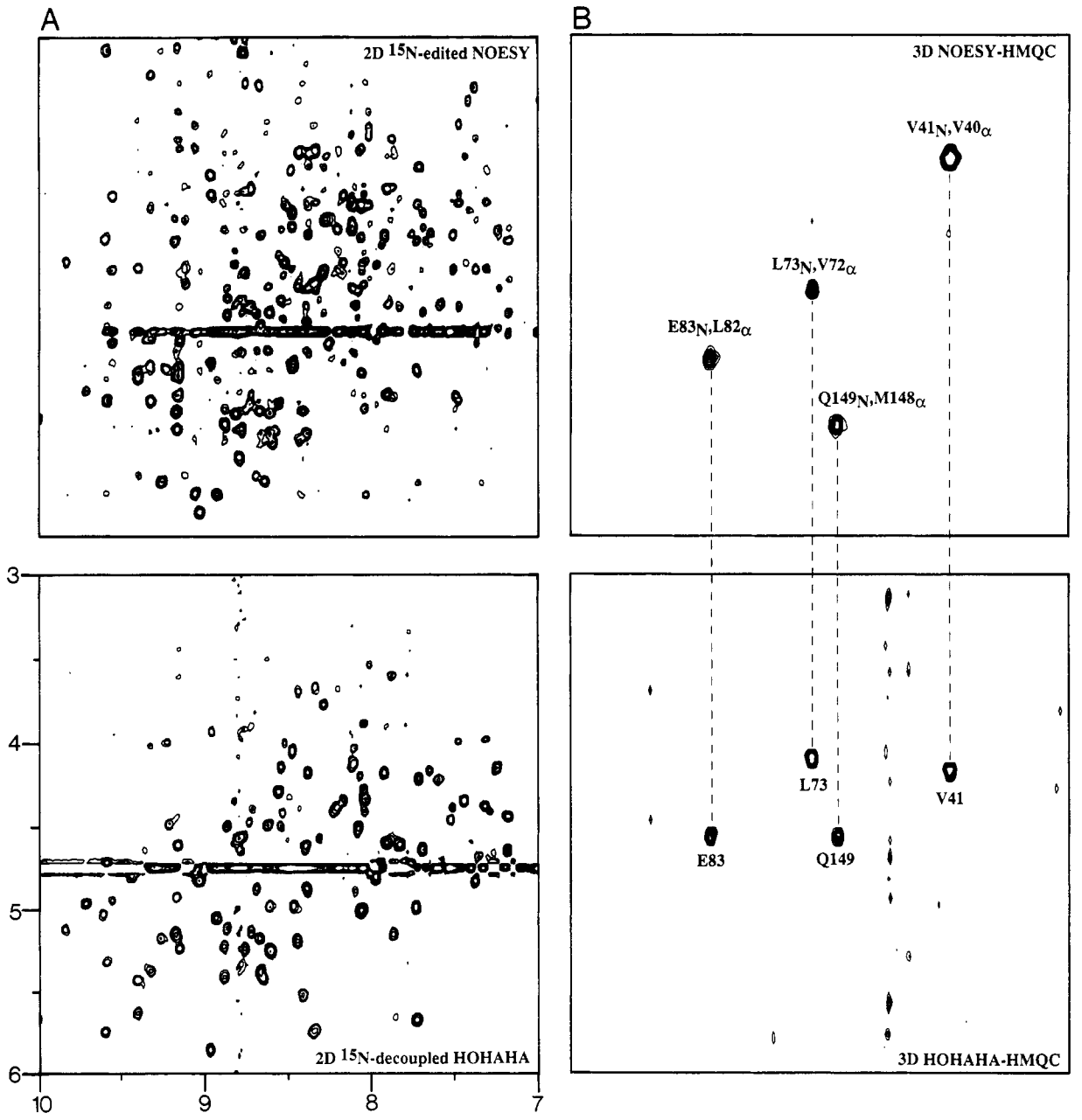


Figure 11. (A) Fingerprint regions of the ^{15}N-edited NOESY (upper) and ^{15}N decoupled HOHAHA (lower) spectra of uniformly ^{15}N labeled interleukin-1β. (B) Corresponding regions of a slice from the 3D heteronuclear NOESY-HMQC (upper) and HOHAHA-HMQC (lower) spectra of the protein under identical conditions (ref 7b).

conformations (60°, 180°, and −60°). In the former case, the coupling and NOE data will be mutually inconsistent, while in the latter, it may not be possible to make an unambiguous distinction between two rotamer conformations. Fortunately, analysis of high-resolution X-ray structures in the protein data bank has shown that 95% of all χ_1 angles lie within ±15° of the staggered rotamer conformations and that there is a very clear correlation between the values of χ_1 and the degree of refinement: the better refined the structures, the closer the χ_1 angles to the ideal staggered rotamer conformations. These results suggest that stereospecific assignments can be obtained for up to 80% of β-methylene protons using this simple approach. In practice, of course, the percentage of stereospecific assignments will be lower due to either spectral overlap or large line widths, preventing the determination of $^3J_{\alpha\beta}$ coupling constants.

A more rigorous approach for stereospecific assignment involves matching the observed $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants, together with approximate distances from the intraresidue $C^\alpha H$–$C^\beta H$ and NH–$C^\beta H$ NOEs and the interresidue $C^\alpha H(i-1)$–$NH(i)$, $C^\alpha H(i)$–$NH(i+1)$, $C^\beta H(i-1)$–$NH(i)$, and $C^\beta H(i)$–$NH(i+1)$ NOEs to theoretical values held in a data base which are derived either for all combinations of ϕ, Ψ, and χ_1 torsion angles (varied by 10°) in a model tripeptide segment or for tripeptide segments taken from high-resolution X-ray structures (*46*, *53*). The data base search is carried out for both possible stereospecific assignments, and in those cases where only one of the two assignments satisfies the information in the data base, the correct stereospecific assignment, together with ranges for the ϕ, ψ, and χ_1 torsion angles, are obtained. A key advantage of this method is that it allows one to obtain much narrower limits for the ϕ, Ψ, and χ_1 torsion angle restraints than would otherwise be possible.

ASSIGNMENT OF LONG RANGE ($|i-j|>5$) NOEs IN PROTEINS

In globular proteins the linear amino acid chain is folded into a tertiary structure such that protons far apart in the

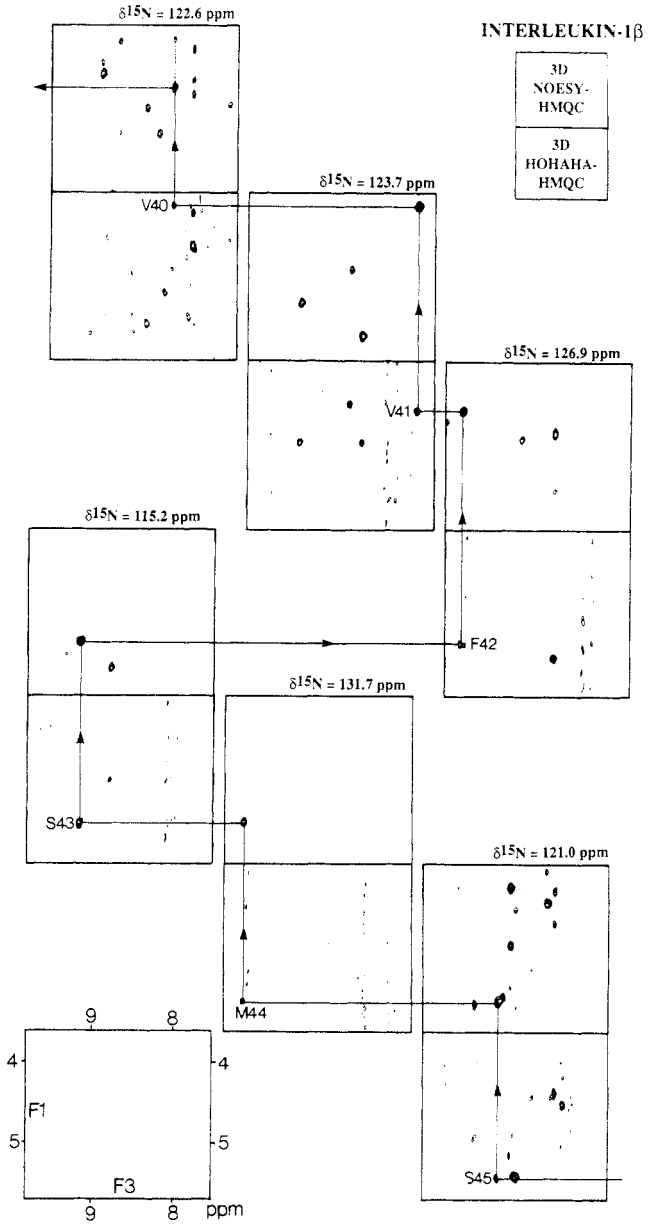


Figure 12. Demonstration of the sequential assignment procedure for the 3D spectra of interleukin-1β. Connectivities between neighboring amino acid spin systems are achieved by "plane hopping" through the 3D cube (ref 7b).

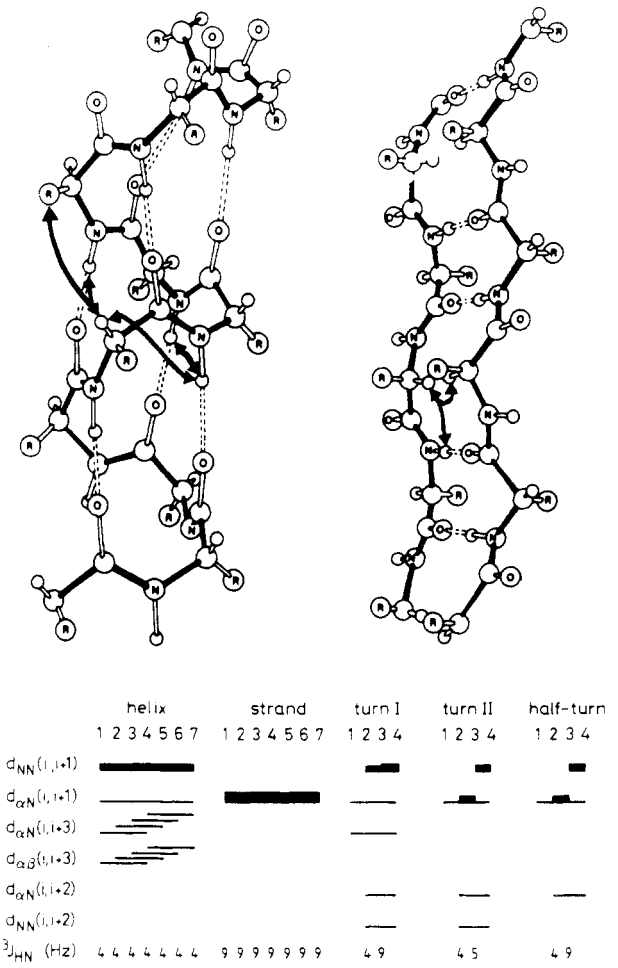


Figure 13. Characteristic patterns of short-range NOEs involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons seen in various regular secondary structure elements. The NOEs are classified as strong, medium, and weak, reflected in the thickness of the lines.

sequence may be close together in space. These protons give rise to tertiary NOEs whose identification is essential for determining the polypeptide fold. Once complete assignments have been made, many such long-range NOEs can be identified in a straightforward manner. It is usually the case, however, that the assignment of a number of long range NOE cross-peaks remains ambiguous due to resonance overlap. In some cases, this ambiguity can be removed by recording additional spectra. Where ambiguities still remain, it is often possible to resolve them by deriving a low-resolution structure on the basis of the available data (i.e. the secondary structure and the assignment of a subset of all the long-range NOEs) either by model building or by distance geometry calculations. This low-resolution structure can then be used to test possible assignments of certain long-range NOEs. For larger proteins it becomes increasingly difficult to assign tertiary NOEs because of the associated overlap problems. In these cases the 3D approach will become a necessity, in particular ^{13}C 3D experiments since they allow editing with respect to particular side chain positions for which the individual NOEs can then be extracted. In Figure 15 all experimental short- and intermediate-range NOE restraints (A) as well as all long-range NOE restraints (B) that were measured from the NOESY spectra are shown as dotted lines superimposed on the framework of the finally determined structure of hirudin, illustrating the dense network of distances throughout the protein core.

TERTIARY STRUCTURE DETERMINATION

Several different approaches can be used to determine the three-dimensional structure of a protein from experimental NMR data. The simplest approach, at least conceptually, is model building. This can be carried out either with real models or by means of interactive molecular graphics. It suffers, however, from the disadvantage that no unbiased measure of the size of the conformational space consistent with the NMR data can be obtained. Consequently, there is no guarantee that the modeled structure is the only one consistent with the experimental data. Further, in this way nothing more than a very low resolution structure can be obtained. Nevertheless, model building can play an important role in the early stages of a structure determination, particularly with respect to resolving ambiguities in the assignments of some of the long-range NOEs.

The main computational methods for generating structures from NMR data comprise as common feature a conformational search to locate the global minimum of a target function that is made up of stereochemical and experimental NMR restraints. The descent to the global minimum region is not a simple straightforward path as the target function is

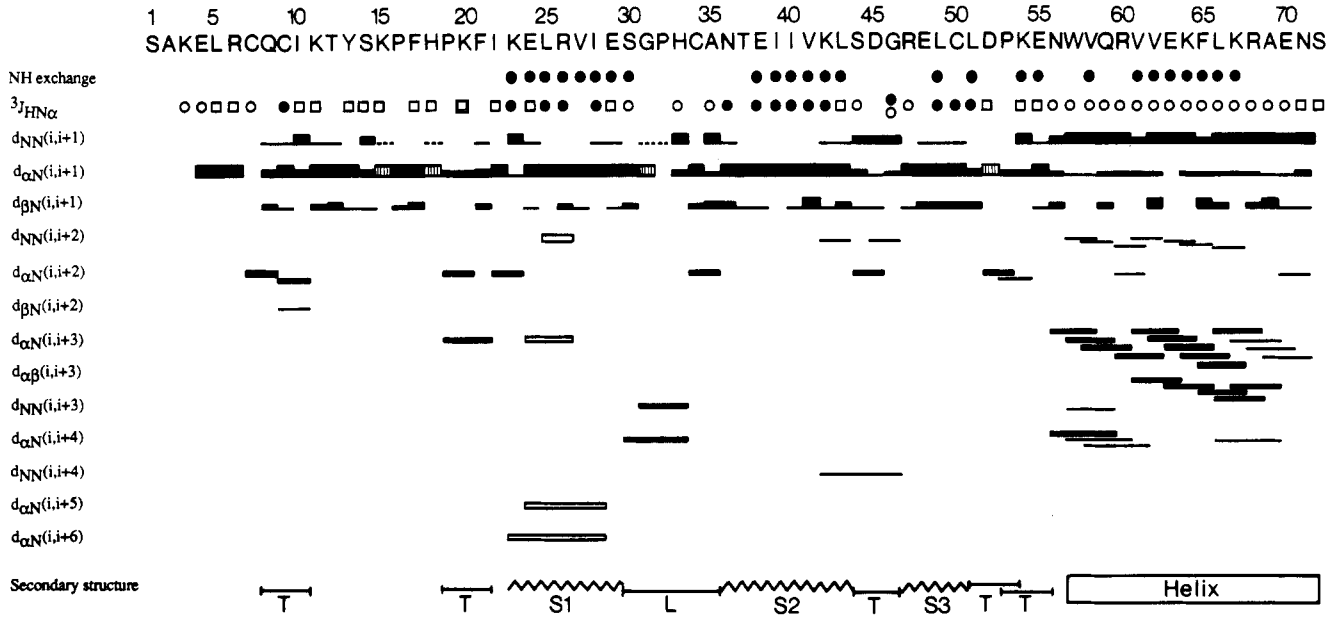


Figure 14. Pattern of short-range NOEs involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons, as well as the $C^{\beta}H$ protons, for interleukin-8. The intensities of the NOEs are indicated by the thickness of the lines. In addition, slowly exchanging backbone amide protons and values of $^3J_{HN\alpha}$ are indicated: $J \leq 7$ Hz (O); 7 Hz $< J <$ 9 Hz (□); $J \geq 9$ Hz (●). The secondary structure deduced from these data is shown at the bottom of the figure (ref 36).

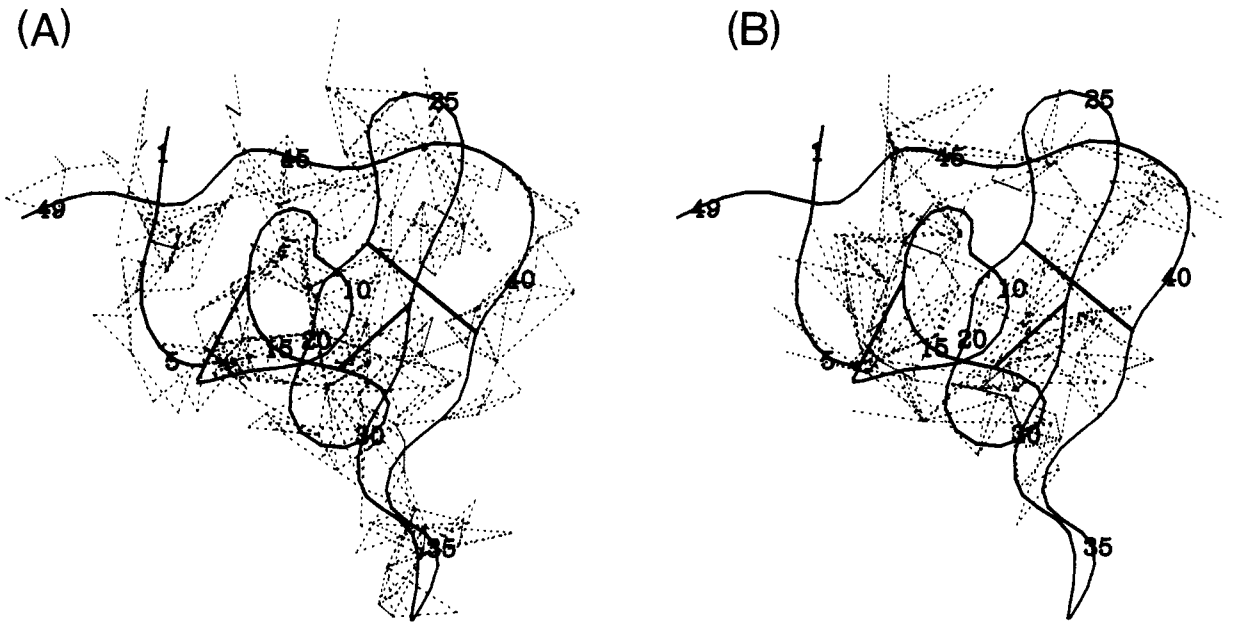


Figure 15. Short-range interresidue ($|i - j| \leq 5$) and intraresidue (A) and long-range interresidue ($|i - j| > 5$) NOE distance restraints (B) shown as dashed lines superimposed on a framework of the final hirudin structure in a smoothed backbone representation.

characterized by many false local minima that have to be avoided or surmounted by all the methods. There are essentially two general classes of methods. The first can be termed real space methods. These include restrained least-squares minimization in torsion angle space with either a variable target function (*9*) or a sequence of ellipsoids of constantly decreasing volume, each of which contains the minimum of the target function (*11*), and restrained molecular dynamics (*10*) and dynamical simulated annealing (*12, 13*) in Cartesian coordinate space. All real space methods require initial structures. These can be (i) random structures with correct covalent geometry; (ii) structures that are very far from the final structure (e.g. a completely extended strand); (iii) structures made up of a completely random array of atoms; and (iv) structures generated by distance space methods. They should not, however, comprise structures derived by model building as this inevitably biases the final outcome. Because these methods operate in real space, great care generally has to be taken to ensure that incorrect folding of the polypeptide chain does not occur. A new real space approach involving the use of dynamical simulated annealing, however, has succeeded in circumventing this problem (*13*). In contrast to the real space methods, the folding problem does not exist in the second class of methods which operates in distance space and is generally referred to as metric matrix distance geometry (*47*). Here the coordinates of the calculated structure are generated by a projection from $N(N-1)/2$ dimensional distance space (where N is the number of atoms) into three-dimensional Cartesian coordinate space by a procedure known as embedding (see ref 48 for a comprehensive review).

A flow chart of the calculational strategy that is generally used to solve protein structures is shown in Figure 16. Since a detailed description of all the various methods would go far

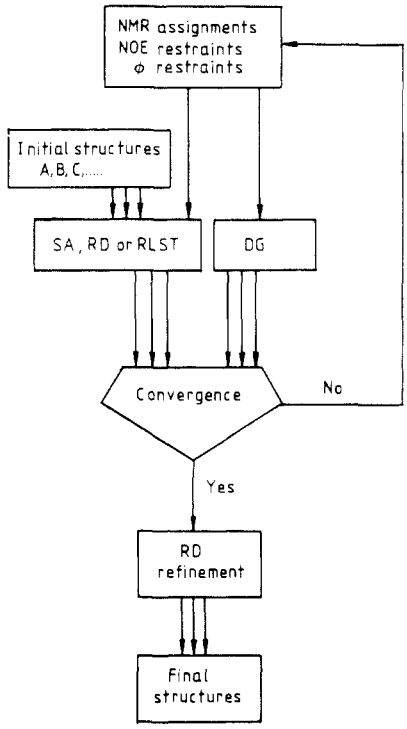


Figure 16. Calculational strategy used to solve three-dimensional structures of macromolecules on the basis of NMR data.

beyond the scope of this review, the interested reader is referred to the references cited above. (A comparison between the different methods is given in ref 49.) All the methods are comparable in convergence power. In general, however, the structures generated by dynamical simulated annealing or refined by restrained molecular dynamics tend to be better in energetic terms than the structures generated by the other methods, particularly with respect to nonbonded contacts and agreement with the experimental NMR data.

In order to assess the uniqueness and the precision of the structures determined by any of the above methods, it is essential to calculate a reasonable number of structures with the same experimental data set, but different starting structures or conditions, and examine their atomic root mean square distribution. If these calculations result in several different folds of the protein while satisfying the experimental restraints, then the data are not sufficient to determine a unique structure and either more data have to be gathered or the structure determination abandoned. If, however, convergence to a single fold is achieved with only small deviations from idealized covalent geometry, exhibiting good nonbonded contacts in addition to satisfying the experimental restraints, then one can be confident that a realistic and accurate picture of the solution structure of the protein has been obtained. The spread observed in the superposition within the family of structures or a plot of the root mean square distribution with respect to the mean allows one to assess the precision associated with different regions of the protein.

EXAMPLES

1. **Hirudin.** Hirudin is a small 65-residue protein from the leech and is the most potent natural inhibitor of coagulation known. It acts by interacting specifically with α-thrombin, thereby preventing the cleavage of fibrinogen. Two recombinant hirudin variants have been examined by NMR: namely wild-type hirudin and the Lys-47 → Glu mutant (*50*). Analysis of the NMR data indicated that hirudin consists of a N-terminal compact domain (residues 1–49) held together by three disulfide bonds and a disordered C-terminal tail

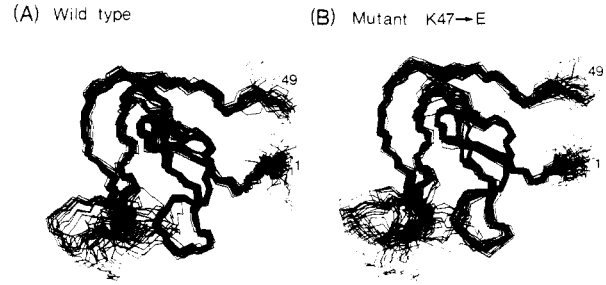


Figure 17. Superposition of the backbone (N, C^{α}, C) atoms of 32 dynamical simulated annealing structures of wild-type hirudin (A) and the Lys-47→Glu mutant (B) for the first 49 amino acids. The wild-type and mutant structures were calculated on the basis of 701 and 677 interproton distance restraints, respectively, 26 ϕ, and 18 χ_1 torsion angle restraints (ref 50).

(residues 50–65). Evidence for the presence of a flexible C-terminal tail was provided by the absence of any intermediate-range or long-range NOEs beyond amino acid 49. Therefore structure calculations were restricted to the N-terminal domain using the hybrid distance geometry–dynamical simulated annealing method. Experimental input data consisted of 701 and 677 approximate interproton distance restraints derived from NOE data for the wild-type and mutant hirudin, respectively, 26 ϕ backbone and 18 χ_1 torsion angle restraints derived from NOE and three-bond coupling constant data, and 8 backbone hydrogen bonds identified on the basis of NOE and amide exchange data. A total of 32 structures were computed for both the wild-type and mutant hirudins (Figure 17). The structure of residues 2–30 and 37–48 constitute the core of the N-terminal domain formed by a triple stranded antiparallel β sheet, and the atomic root mean square difference between the individual structures and the mean structure is ~0.7 Å for the backbone atoms and ~1 Å for all atoms. The orientation of the exposed finger of antiparallel β sheet (residues 31–36) with respect to the core could not be determined as no long range NOEs were observed between the exposed finger and the core. This is easily appreciated from Figure 17 since the structures in that region exhibit a very large spread. Locally, however, the polypeptide fold of residues 31–36 is reasonably well-defined.

The first five residues form an irregular strand which leads into a loop closed off at its base by the disulfide bridge between Cys-6 and Cys-14. This is followed by a mini-antiparallel β sheet formed by residues 14–16 (strand I) and 21 and 22 (strand I′) connected by a type II turn. This β sheet is distorted by a β bulge at Cys-16. Strand I′ leads into a second antiparallel β sheet formed by residues 27–31 (strand II) and 36–40 (strand II′) connected by a β turn (residues 32–35). Additionally, residues 10 and 11 exhibit features of a β bulge with the amide of Gly-10 and the carbonyl oxygen atom of Glu-11 hydrogen bonded to the carbonyl and amide groups, respectively, of Cys-28. Finally, strand II′ leads into an irregular strand which folds back onto the protein such that residue 47 (Lys in the wild type, Glu in the mutant) is in close proximity to residues in the loop closed off by the disulfide bridge between Cys-6 and Cys-14. Not only the backbone but also many of the side chain conformations are well-defined, especially those in the interior of the protein.

A superposition of the core (residues 1–30 and 37–49) of the restrained minimized mean structures of the wild-type and mutant hirudin provides a good representation of the differences between the two structures (Figure 18A). Regions of noticeable difference can be identified where the atomic root mean square difference between the two mean structures is larger than the atomic root mean square distribution of the individual structures about their respective means. This analysis indicates the presence of clear differences for the

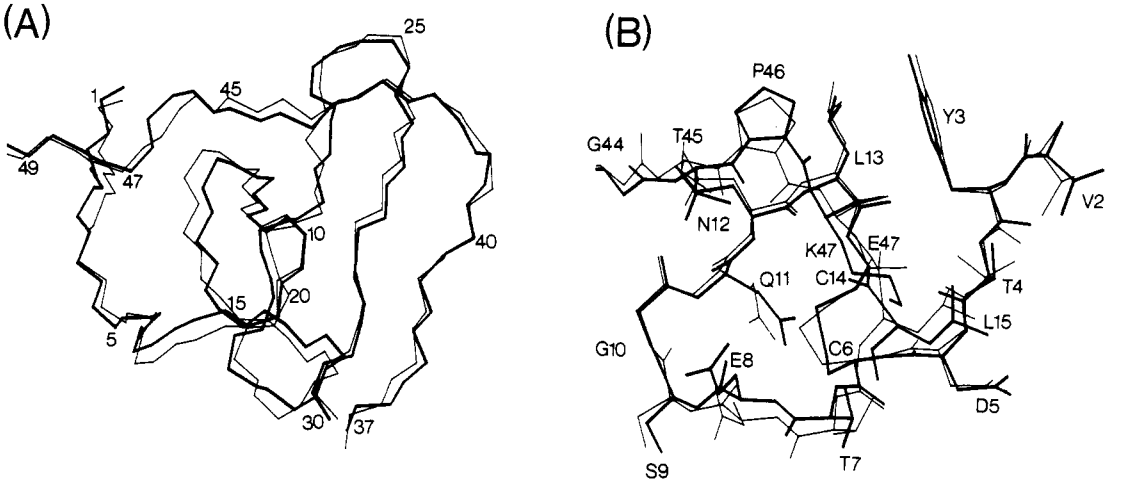


Figure 18. (A) Superposition of the core (residues 1–30 and 37–49) of the restrained minimized mean structure of wild-type (thick lines) and Lys-47→Glu mutant (thin lines) hirudin. (B) View of all atoms around the site of mutation (ref 50).

backbone atoms of residues 3, 5, 8, 11–15, 22, 26, and 27. In the mutant, the backbone atoms of residues 3, 4, and 11–15 are slightly closer to that of residues 45–47 than in the wild-type, a change which can be rationalized in terms of the shorter length of the Glu side chain relative to that of Lys. Concomitantly, the backbone of residue 8 appears to be pushed away in the wild type structure. Residues 22, 26, and 27 also move in the same direction as residues 11–15. This is secondary to the perturbation of residues 11–15 and can be attributed to the presence of numerous contacts between residues 8–11 on the one hand and 28–30 on the other, including hydrogen bonds between Cys-28 and Gly-10 and between Cys-28 and Glu-11. There seems to be no significant differences, however, with respect to the side chain positions within the errors of the coordinates, even within the immediate vicinity of residue 47 (Figure 18B).

At this point it may be worth mentioning that NMR can have two important applications with respect to genetically engineered proteins. The structural studies on hirudin were first initiated by using natural protein extracted from the whole body of leeches (*51*). All further work was subsequently carried out on recombinant products, either comprising the wild-type sequence or mutants thereof (*50*). It was a fast and easy task to assess the structural identity of the recombinant product, since only a comparison of the appropriate 2D spectra had to be carried out. These spectra can be regarded as a fingerprint of a particular protein and, hence, a reflection of the 3D structure in solution. Simple overlay often allows one to ascertain whether the recombinant product is folded indistinguishably from the natural counterpart. In the same way, NMR can establish whether the structure of a mutant is essentially unchanged from the wild type one, since very similar spectra will arise from closely related structures. Analyses of this kind can be carried out without a full assignment of the spectra or a complete structure calculation, thus providing a useful tool for guiding genetic engineering projects.

2. Cellulose Binding Domain of Cellobiohydrolase. Cellulases are enzymes involved in plant cell wall degradation that exhibit a common domain structure consisting of a catalytic domain (~400–500 amino acids), a highly conserved (~70% sequence identity) terminal domain (~40 amino acids) that is located either at the C terminus or N terminus, and a heavily glycosylated linker region (~30 amino acids), which connects the two domains. The domain architecture of the cellulases as well as the observation that the domains of CBH I retain their respective activities after cleavage suggested a dual approach to the problem of obtaining a three-dimensional structure of this cellulase, involving the

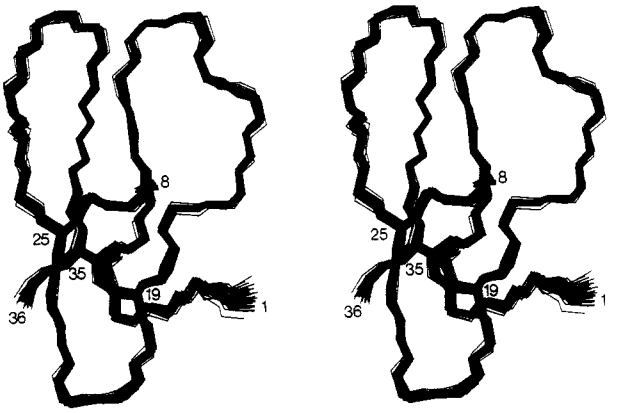


Figure 19. Superposition of the backbone (N, C^{α}, C) atoms of 41 dynamical simulated annealing structures of the C-terminal domain of CBH I (ref 53).

application of both NMR and X-ray crystallography. To date there has been no success in the crystallization of an intact cellulase, whereas crystals of the catalytic domain of CBH II have been obtained (*52*), opening the possibility for determining its X-ray structure. The cellulose binding domain being a very small polypeptide, on the other hand, seemed ideally suited for a structure determination by NMR. Because of its small size it was possible to chemically synthesize the 36 amino acid domain and it was established that the biological properties were identical to those of the cleavage product. The subsequent NMR structure determination (*53*) was therefore carried out on the synthetic product and a large number of stereospecific assignments obtained by using the data base approach resulted in numerous interproton distance and torsion angle restraints. Thus an exceptionally large experimental data set consisting of 554 interproton distance restraints, 33ϕ, 24Ψ, and 25χ_1 torsion angle restraints, and 42 hydrogen bonding restraints was used in the structure determination. The converged set of 41 structures represents the best quality NMR structure to date, exhibiting a root mean square difference between the individual structures and the mean coordinate positions of 0.33 Å for the backbone atoms and 0.52Å for all atoms. It was possible to determine the pairing of the two disulfide bridges, which previously was unknown. A backbone trace of the CBH I structure is shown in Figure 19 and several regions including the amino acid side chains are displayed in Figure 20. The protein has a wedgelike shape with an amphiphilic character, one face being predominantly hydrophobic and the other mainly hydrophilic. As can be readily appreciated from Figures 19 and 20, the quality

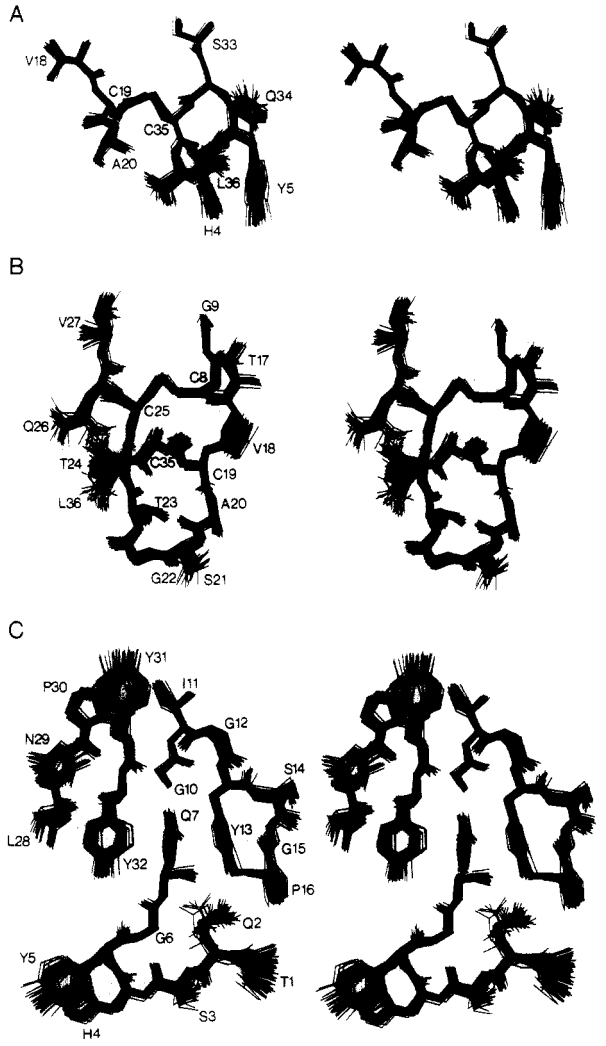


Figure 20. Best fit superpositions (41 structures) of all atoms for three selected regions of the C-terminal domain of CBH I illustrating the excellent definition of side chain positions (ref 53).

of this structure is comparable to a crystal structure at approximately 2–2.5 Å resolution, and such structures permit the detailed analysis of side chain–side chain interactions and other possibly interesting structural features.

PERSPECTIVE AND CONCLUDING REMARKS

It should be clear from the above discussion that NMR now stands side by side with X-ray crystallography as a powerful method for three-dimensional structure determination. What are the limitations of this approach? At present it is limited to proteins of molecular weight ≤20 000. Indeed the largest proteins whose three-dimensional structures have been determined to date are plastocyanin (99 residues (*54*)), the globular domain of histone H5 (79 residues (*55*), α-amylase inhibitor (74 residues (*56*)), and interleukin-8 (a dimer of 72 residues per monomer (*57*)). Virtually complete assignments, however, have been made for a variety of larger systems, in particular hen egg white lysozyme (129 residues (*58*)) and the lac repressor headpiece–operator complex (molecular weight ~15 000 (*59*)), *Staphylococcal* nuclease (148 residues (*60*)), and interleukin-1β (153 residues (*61*)). Further development of novel techniques based on multidimensional NMR in combination with isotopic labeling and the introduction of yet more powerful magnets may make it possible to extend the molecular weight range up to proteins of molecular weight ~40 000 in the future. This, however, will probably present a fundamental limit as the large line widths of such proteins significantly reduce even the sensitivity of ^{1}H-detected heteronuclear correlation experiments.

At this point it is appropriate to add a word of caution concerning the practical limits of structure determination by NMR. It is not always the size or the number of residues in a particular protein that determines the feasibility of an NMR structure determination. Other factors play equally important roles. For example, the protein should be soluble up to millimolar concentrations, nonaggregating, and preferably stable up to at least 40 °C, particularly for large proteins. A further consideration is the chemical shift dispersion of the ^{1}H NMR spectrum. This depends to a large extent on the structure of the protein under investigation. Proteins that are made up only of α-helices, loops and turns, invariably exhibit fairly poor proton chemical shift dispersion, while the chemical shift dispersion in β-sheet proteins is usually very good. To a certain degree such problems associated with chemical shift degeneracy can be overcome by heteronuclear 3D experiments.

Another potential problem may arise from the fact that different regions or domains of a protein may be well-defined and therefore amenable to a NMR structure determination, while other parts of the same protein may not. This will be reflected in the absence of long-range tertiary NOEs for the ill-defined regions and leads to the inability to position these regions with respect to the rest of the protein satisfactorily (e.g. the case of hirudin discussed above). We therefore believe that it is necessary to calculate a reasonable number of structures (ca. 20) with the same experimental data set in order to obtain a good representation of the NMR structure. Only by analyzing such a family of structures can the local and global definition of a structure be assessed.

X-ray crystallography, of course, also has its limitations, the most obvious being the requirement for a protein to crystallize. Thus suitable crystals that diffract to high resolution have to be grown and a successful search for heavy atom derivatives to solve the phase problem is necessary. Therefore not every protein will be amenable to both NMR and X-ray crystallography. In those cases where this is feasable, the information afforded by NMR and crystallography is clearly complementary and may lead to a deeper understanding of the differences between the solution and crystalline state of the protein.

Finally it should be stressed that, in addition to being able to determine three-dimensional structures of proteins, NMR has the potential to address other questions, in particular those concerning the dynamics of the system. This opens the possibility that a whole variety of different NMR studies can be initiated on the basis of an NMR structure, such as the investigation of the dynamics of conformational changes upon ligand binding, unfolding kinetics, conformational equilibria between different conformational states, fast internal dynamics on the nanosecond time scale and below, and slow internal motions on the second and millisecond time scales. The results obtained from these kinetic studies can then be interpreted in the light of the previously determined structure, thus bringing together structure and dynamics of proteins in a unified picture.

ACKNOWLEDGMENT

The work in the authors' laboratory was supported in part by the AIDS targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health.

LITERATURE CITED

(1) Jeener, J., Ampere International Summer School, Basko Polje, Yugoslavia, 1971, unpublished lecture.
(2) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys.* **1976**, *64*, 2229.
(3) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546.
(4) Oschkinat, H.; Griesinger, C.; Kraulis, P. J.; Sørensen, O. W.; Ernst, R. R.; Gronenborn, A. M.; Clore, G. M. *Nature (London)* **1988**, *332*,

374–377. Oschkinat, H.; Cieslar, C.; Holak, T. A.; Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1989**, *83*, 450.
(5) Vuister, G. W.; Boelens, R.; Kaptein, R. *J. Magn. Reson.* **1988**, *80*, 176–185.
(6) Fesik, S. W.; Zuiderweg, E. R. P. *J. Magn. Reson.* **1988**, *78*, 588–593. Zuiderweg, E. R. P.; Fesik, S. W. *Biochemistry* **1989**, *28*, 2387.
(7) Marion, D.; Kay, L. E.; Sparks, S. W.; Torchia, D. A.; Bax, A. *J. Am. Chem. Soc.* **1989**, *111*, 1515. Marion, D.; Driscoll, P. C.; Kay, L. E.; Wingfield, P. T.; Bax, A.; Gronenborn, A. M.; Clore, G. M. *Biochemistry* **1989**, *28*, 6150.
(8) Havel, T. F.; Kuntz, I. D.; Crippen, G. M. *Math. Biol.* **1983**, *45*, 665. Havel, T. F.; Wüthrich, K. *Bull. Math. Biol.* **1984**, *46*, 673.
(9) Braun, W.; Go, N. *J. Mol. Biol.* **1985**, *186*, 611.
(10) Clore, G. M.; Gronenborn, A. M.; Brünger, A. T.; Karplus, M. *J. Mol. Biol.* **1985**, *186*, 435. Kaptein, R.; Zuiderweg, E. R. P.; Scheek, R. M.; Boelens, R.; van Gunsteren, W. F. *J. Mol. Biol.* **1985**, *182*, 179. Clore, G. M.; Brünger, A. T.; Karplus, M.; Gronenborn, A. M. *J. Mol. Biol.* **1986**, *191*, 523.
(11) Billeter, M.; Havel, T. F.; Wüthrich, K. *J. Comput. Chem.* **1987**, *8*, 132.
(12) Nilges, M.; Gronenborn, A. M.; Clore, G. M. *FEBS Lett.* **1988**, *229*, 317. Nilges, M.; Gronenborn, A. M.; Brünger, A. T.; Clore, G. M. *Protein Eng.* **1988**, *2*, 27.
(13) Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1988**, *239*, 129.
(14) Litharge, O.; Cornelius, C. W.; Buchanan, B. G.; Jardetzky, O. *Proteins: Struct., Funct. Genet.* **1987**, *2*, 340.
(15) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*: Clarendon Press: Oxford, 1986.
(16) Davis, D. G.; Bax, A. *J. Am. Chem. Soc.* **1985**, *107*, 2821. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355. Bax, A. In *Methods in Enzymology*; James, T. L., Oppenheimer, N., Eds.; Academic Press: New York, 1988; Vol. 176, p 151.
(17) Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95. Macura, S.; Huang, Y.; Suter, D.; Ernst, R. R. *J. Magn. Reson.* **1981**, *43*, 259.
(18) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Magn. Reson.* **1989**, *84*, 14.
(19) Overhauser, A. *Phys. Rev.* **1953**, *89*, 689. Overhauser, A. *Phys. Rev.* **1953**, *92*, 411.
(20) Solomon, I. *Phys. Rev.* **1955**, *99*, 559.
(21) Noggle, J. H.; Schirmer, R. E. *The Nuclear Overhauser Effect—Chemical Applications*; Academic Press: New York, 1971.
(22) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; J. Wiley: New York, 1986.
(23) Rance, M.; Sørenson, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479.
(24) Marion, D.; Bax, A. *J. Magn. Reson.* **1988**, *79*, 352.
(25) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521.
(26) Bax, A.; Kay, L. E.; Sparks, S. W.; Torchia, D. L. *J. Am. Chem. Soc.* **1989**, *111*, 408.
(27) Mueller, L. *J. Am. Chem. Soc.* **1979**, *101*, 4481.
(28) Redfield, A. G. *Chem. Phys. Lett.* **1983**, *96*, 537. Griffey, R. H.; Redfield, A. G. *Q. Rev. Biophys.* **1987**, *19*, 51.
(29) Bax, A.; Griffey, R. H.; Hawkins, B. L. *J. Am. Chem. Soc.* **1983**, *105*, 7188.
(30) Bolton, P. H.; Bodenhausen, G. *Chem. Phys. Lett.* **1982**, *89*, 139.
(31) Sklenar, V.; Bax, A. *J. Magn. reson.* **1987**, *71*, 379.
(32) Gronenborn, A. M.; Bax, A.; Wingfield, P. T.; Clore, G. M. *FEBS Lett.* **1989**, *243*, 93. Gronenborn, A. M.; Wingfield, P. T.; Clore, G. M. *Biochemistry* **1989**, *28*, 5081.
(33) Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2093. Bax, A.; Marion, D. *J. Magn. reson.* **1988**, *78*, 186.
(34) Clore, G. M.; Bax, A.; Wingfield, P. T.; Gronenborn, A. M. *FEBS Lett.* **1988**, *238*, 17.
(35) Wüthrich, K., Wider, G., Wagner, G. and Braun, W. *J. Mol. Biol.* **1982**, *155*, 311. Billeter, M.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1982**, *155*, 321. Wagner, G.; Wüthrich, K. *J. Mol. Biol.* **1982**, *155*, 347. Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715.
(36) Clore, G. M.; Appella, E.; Yamada, M.; Matsushima, K.; Gronenborn, A. M. *J. Biol. Chem.*, in press.
(37) Wagner, G.; Wüthrich, K. *J. Magn. Reson.* **1979**, *33*, 675. Dobson, C. M.; Olejniczak, E. T.; Poulsen, F. M.; Ratcliffe, R. G. *J. Magn. Reson.* **1982**, *48*, 97. Keepers, J. W.; James, T. L. *J. Magn. Reson.* **1984**, *57*, 404. Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1985**, *61*, 158.
(38) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1988**, *79*, 493.
(39) Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1989**, *84*, 398.
(40) Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870. Pardi, A.; Billeter, M.; Wüthrich, K. *J. Mol. Biol.* **1984**, *180*, 741.
(41) Neuhaus, D.; Wagner, G.; Vasák, M.; Kägi, J. H. R.; Wüthrich, K. *Eur. J. Biochem.* **1985**, *151*, 257.
(42) Kay, L. E.; Bax, A. *J. Magn. Reson.*, in press. Forman-Kay, J. D.; Gronenborn, A. M.; Kay, L. E.; Wingfield, P. T.; Clore, G. M. *Biochemistry*, in press.
(43) Mueller, L. *J. Magn. Reson.* **1987**, *203*, 251.
(44) Oschkinat, H.; Pastore, A.; Pfändler, P.; Bodenhausen, G. *J. Magn. Reson.* **1986**, *69*, 559. Oschkinat, H.; Clore, G. M.; Nilges, M.; Gronenborn, A. M. *J. Magn. Reson.* **1987**, *75*, 534.
(45) Driscoll, P. C.; Clore, G. M.; Beress, L.; Gronenborn, A. M. *Biochemistry* **1989**, *28*, 2178.
(46) Nilges, M.; Clore, G. M.; Gronenborn, A. M. *Biopolymers*, in press.
(47) Crippen, G. M. *J. Comp. Phys.* **1977**, *24*, 96.
(48) Crippen, G. M.; Havel, T. F. *Distance Geometry and Molecular Conformation*; John Wiley: New York, 1988.
(49) Clore, G. M.; Gronenborn, A. M. *CRC Crit. Rev. Biochem.* **1989**, *24*, 479.
(50) Folkers, P. J. M.; Clore, G. M.; Driscoll, P. C.; Dodt, J.; Köhler, S.; Gronenborn, A. M. *Biochemistry* **1988**, *28*, 2601.
(51) Clore, G. M.; Sukumaran, D. K.; Nilges, M.; Zarbock, J.; Gronenborn, A. M. *EMBO J.* **1987**, *6*, 529.
(52) Bergfors, T.; Rouvinen, J.; Lehtovaara, P.; Caldentey, X.; Tomme, P.; Claeyssens, M.; Pettersson, G.; Teeri, T.; Knowles, J.; Jones, T. A. *J. Mol. Biol.* **1989**, *209*, 167.
(53) Kraulis, P. J.; Clore, G. M.; Nilges, M.; Jones, T. A.; Pettersson, G.; Knowles, J.; Gronenborn, A. M. *Biochemistry* **1989**, *28*, 7241.
(54) Moore, J. M.; Case, D. A.; Chazin, W. J.; Gippert, G. P.; Havel, T. F.; Powls, R.; Wright, P. E. *Science* **1988**, *240*, 314.
(55) Clore, G. M.; Gronenborn, A. M.; Nilges, M.; Sukumaran, D. K.; Zarbock, J. *EMBO J.* **1987**, *6*, 1833.
(56) Kline, A. D.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1986**, *189*, 377. Kline, A. D.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1988**, *204*, 675.
(57) Clore, G. M.; Appella, E.; Yamada, M.; Matsushima, K.; Gronenborn, A. M., submitted for publication.
(58) Redfield, C.; Dobson, C. M. *Biochemistry* **1988**, *27*, 122.
(59) Boelens, R.; Scheek, R. M.; Lamerichs, R. M. J. N.; de Vlieg, J.; van Boom, J. H.; Kaptein, R. *J. Mol. Biol.* **1987**, *193*, 213.
(60) Torchia, D. A.; Sparks, S. W.; Bax, A. *Biochemistry* **1988**, *27*, 5135.
(61) Driscoll, P. C.; Clore, G. M.; Marion, D.; Wingfield, P. T.; Gronenborn, A. M. *Biochemistry*, in press.